

2/13 #6



OIPE

RAW SEQUENCE LISTING

DATE: 02/19/2002

PATENT APPLICATION: US/09/881,752A

TIME: 14:40:02

Input Set : A:\06132.041002.SEQLIST.TXT

Output Set: N:\CRF3\02192002\I881752A.raw

P.S
ENTERED

4 <110> APPLICANT: Kleanthous, Harold
 5 Al-Garawi, Amal
 6 Miller, Charles
 7 Tomb, Jean-Francois
 8 Oomen, Raymond P.
 10 <120> TITLE OF INVENTION: Identification of Polynucleotides
 11 Encoding Novel Helicobacter Polypeptides in the Helicobacter
 12 Genome
 14 <130> FILE REFERENCE: 06132/041002
 16 <140> CURRENT APPLICATION NUMBER: US 09/881,752A
 17 <141> CURRENT FILING DATE: 2001-06-15
 19 <150> PRIOR APPLICATION NUMBER: US 08/833,457
 20 <151> PRIOR FILING DATE: 1997-04-01
 22 <160> NUMBER OF SEQ ID NOS: 370
 24 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 26 <210> SEQ ID NO: 1
 27 <211> LENGTH: 265
 28 <212> TYPE: DNA
 29 <213> ORGANISM: Helicobacter pylori
 31 <220> FEATURE:
 32 <221> NAME/KEY: CDS
 33 <222> LOCATION: (51)...(212)
 35 <400> SEQUENCE: 1
 36 ttttttagtt tgtttttgag tataatccta cgaaaatttt aaggaacggc atg gag 56
 37 Met Glu
 38 1
 40 ttt ttg gga ctg att tta agt ctg gcc gct att ttg ata gcg ttt aaa 104
 41 Phe Leu Gly Leu Ile Leu Ser Leu Ala Ala Ile Leu Ile Ala Phe Lys
 42 5 10 15
 44 aag cct gaa aaa gaa aat tgg gcg ttt ggg att ttg atg gtg gtg tgg 152
 45 Lys Pro Glu Lys Glu Asn Trp Ala Phe Gly Ile Leu Met Val Val Trp
 46 20 25 30
 48 tta gtg gag ctt att att ttt ata gcc cac agc tct agc gtt ttg cct 200
 49 Leu Val Glu Leu Ile Ile Phe Ile Ala His Ser Ser Ser Val Leu Pro
 50 35 40 45 50
 52 aac atg aat cta taagggggat gcatggataa agaaaccgga ttttacaacc 252
 53 Asn Met Asn Leu
 54 ttttttcttt ggc 265
 58 <210> SEQ ID NO: 2
 59 <211> LENGTH: 54
 60 <212> TYPE: PRT
 61 <213> ORGANISM: Helicobacter pylori
 63 <400> SEQUENCE: 2

RAW SEQUENCE LISTING

DATE: 02/19/2002

PATENT APPLICATION: US/09/881,752A

TIME: 14:40:02

Input Set : A:\06132.041002.SEQLIST.TXT

Output Set: N:\CRF3\02192002\I881752A.raw

```

64 Met Glu Phe Leu Gly Leu Ile Leu Ser Leu Ala Ala Ile Leu Ile Ala
65 1 5 10 15
66 Phe Lys Lys Pro Glu Lys Glu Asn Trp Ala Phe Gly Ile Leu Met Val
67 20 25 30
68 Val Trp Leu Val Glu Leu Ile Ile Phe Ile Ala His Ser Ser Ser Val
69 35 40 45
70 Leu Pro Asn Met Asn Leu
71 50
74 <210> SEQ ID NO: 3
75 <211> LENGTH: 670
76 <212> TYPE: DNA
77 <213> ORGANISM: Helicobacter pylori
79 <220> FEATURE:
80 <221> NAME/KEY: CDS
81 <222> LOCATION: (51)...(617)
83 <400> SEQUENCE: 3
84 cccatagacg acaaaatcaa gcggttttat caaaacccaaa aaactttaga atg aaa 56
85 Met Lys
86 1
88 aaa att gct ttc att ttg gct tta tgg gtg ggc ttg tta ggg gcg ttt 104
89 Lys Ile Ala Phe Ile Leu Ala Leu Trp Val Gly Leu Leu Gly Ala Phe
90 5 10 15
92 gag cct aaa aaa agt cat att tat ttt ggg gct atg gtg ggt tta gct 152
93 Glu Pro Lys Lys Ser His Ile Tyr Phe Gly Ala Met Val Gly Leu Ala
94 20 25 30
96 cct att aaa ata acc cca aaa ccg gct agt gat tct tct tat acg gct 200
97 Pro Ile Lys Ile Thr Pro Lys Pro Ala Ser Asp Ser Ser Tyr Thr Ala
98 35 40 45 50
100 ttt tta tgg ggg gct aaa gga ggg tat caa ttc gct ttt ttt aaa gct 248
101 Phe Leu Trp Gly Ala Lys Gly Gly Tyr Gln Phe Ala Phe Phe Lys Ala
102 55 60 65
104 cta gcg tta agg ggt gaa ttt tcc tac ctt atg gca atc aaa ccc acc 296
105 Leu Ala Leu Arg Gly Glu Phe Ser Tyr Leu Met Ala Ile Lys Pro Thr
106 70 75 80
108 gca ctg cac acg att aac act tct tta ttg agc tta aat att gat gtg 344
109 Ala Leu His Thr Ile Asn Thr Ser Leu Leu Ser Leu Asn Ile Asp Val
110 85 90 95
112 tta agc gat ttt tac act tac aaa aaa tac agc ttt ggg gtg tat ggg 392
113 Leu Ser Asp Phe Tyr Thr Tyr Lys Lys Tyr Ser Phe Gly Val Tyr Gly
114 100 105 110
116 ggg ctt ggg ata ggg tat ttt tat caa agc aac cat tta ggc atg aaa 440
117 Gly Leu Gly Ile Gly Tyr Phe Tyr Gln Ser Asn His Leu Gly Met Lys
118 115 120 125 130
120 aat agt tcg ttt atg ggt tat aac ggc ttg ttt aat gtg ggg ctt ggc 488
121 Asn Ser Ser Phe Met Gly Tyr Asn Gly Leu Phe Asn Val Gly Leu Gly
122 135 140 145
124 agc acg atc gat cgc cac cac cgc ata gag ctt ggg gct aaa atc cct 536
125 Ser Thr Ile Asp Arg His His Arg Ile Glu Leu Gly Ala Lys Ile Pro
126 150 155 160

```

RAW SEQUENCE LISTING

DATE: 02/19/2002

PATENT APPLICATION: US/09/881,752A

TIME: 14:40:02

Input Set : A:\06132.041002.SEQLIST.TXT

Output Set: N:\CRF3\02192002\I881752A.raw

```

128 ttt tca aag act aga aat tct ttt aaa aat cct tat ttt tta gag agc 584
129 Phe Ser Lys Thr Arg Asn Ser Phe Lys Asn Pro Tyr Phe Leu Glu Ser
130      165      170      175
132 gtt ttt atc cat gcg act tat agc tat atg ttt taagagagaa tagcctatta 637
133 Val Phe Ile His Ala Thr Tyr Ser Tyr Met Phe
134      180      185
136 gtggtcgtta tcaataagat aagatcotta atg 670
138 <210> SEQ ID NO: 4
139 <211> LENGTH: 189
140 <212> TYPE: PRT
141 <213> ORGANISM: Helicobacter pylori
143 <400> SEQUENCE: 4
144 Met Lys Lys Ile Ala Phe Ile Leu Ala Leu Trp Val Gly Leu Leu Gly
145 1      5      10      15
146 Ala Phe Glu Pro Lys Lys Ser His Ile Tyr Phe Gly Ala Met Val Gly
147      20      25      30
148 Leu Ala Pro Ile Lys Ile Thr Pro Lys Pro Ala Ser Asp Ser Ser Tyr
149      35      40      45
150 Thr Ala Phe Leu Trp Gly Ala Lys Gly Gly Tyr Gln Phe Ala Phe Phe
151      50      55      60
152 Lys Ala Leu Ala Leu Arg Gly Glu Phe Ser Tyr Leu Met Ala Ile Lys
153 65      70      75      80
154 Pro Thr Ala Leu His Thr Ile Asn Thr Ser Leu Leu Ser Leu Asn Ile
155      85      90      95
156 Asp Val Leu Ser Asp Phe Tyr Thr Tyr Lys Lys Tyr Ser Phe Gly Val
157      100     105     110
158 Tyr Gly Gly Leu Gly Ile Gly Tyr Phe Tyr Gln Ser Asn His Leu Gly
159      115     120     125
160 Met Lys Asn Ser Ser Phe Met Gly Tyr Asn Gly Leu Phe Asn Val Gly
161      130     135     140
162 Leu Gly Ser Thr Ile Asp Arg His His Arg Ile Glu Leu Gly Ala Lys
163 145     150     155     160
164 Ile Pro Phe Ser Lys Thr Arg Asn Ser Phe Lys Asn Pro Tyr Phe Leu
165      165     170     175
166 Glu Ser Val Phe Ile His Ala Thr Tyr Ser Tyr Met Phe
167      180     185
170 <210> SEQ ID NO: 5
171 <211> LENGTH: 434
172 <212> TYPE: DNA
173 <213> ORGANISM: Helicobacter pylori
175 <220> FEATURE:
176 <221> NAME/KEY: CDS
177 <222> LOCATION: (51)...(380)
179 <400> SEQUENCE: 5
180 agcgtgaaaa aaattgagtt gaatcaaaac ctgcattaag gattaaaaga atg ctc 56
181      Met Leu
182      1
184 aaa aaa agt ttg tta ttg ctt gtt ttt tta gtc tta cag ctt agc ggc 104
185 Lys Lys Ser Leu Leu Leu Val Phe Leu Val Leu Gln Leu Ser Gly

```

RAW SEQUENCE LISTING

DATE: 02/19/2002

PATENT APPLICATION: US/09/881,752A

TIME: 14:40:02

Input Set : A:\06132.041002.SEQLIST.TXT

Output Set: N:\CRF3\02192002\I881752A.raw

```

186          5          10          15
188 gct gaa gaa aac aat caa gcc cca aaa aac acg ccc cct gaa tta aac 152
189 Ala Glu Glu Asn Asn Gln Ala Pro Lys Asn Thr Pro Pro Glu Leu Asn
190          20          25          30
192 ccc gct aac gct aag ggc gcg cca aac tct aac acc cag atc acc cct 200
193 Pro Ala Asn Ala Lys Gly Ala Pro Asn Ser Asn Thr Gln Ile Thr Pro
194 35          40          45          50
196 aaa aac gat aac tct aac ctg tta gac aaa tta ggt tcg cct gaa aac 248
197 Lys Asn Asp Asn Ser Asn Leu Leu Asp Lys Leu Gly Ser Pro Glu Asn
198          55          60          65
200 gct caa acc gag ctt tct gcc ggt att gat ttg gct aaa aag ggc gat 296
201 Ala Gln Thr Glu Leu Ser Ala Gly Ile Asp Leu Ala Lys Lys Gly Asp
202          70          75          80
204 tat caa ggg gct ttc aag ctt ttt tcc caa tcg tgc gat aat ggt aat 344
205 Tyr Gln Gly Ala Phe Lys Leu Phe Ser Gln Ser Cys Asp Asn Gly Asn
206          85          90          95
208 gcg gcc ggg tgt ttt gca agt ggg ggc gat gta tgc taatggggta 390
209 Ala Ala Gly Cys Phe Ala Ser Gly Gly Asp Val Cys
210 100          105          110
212 gggatccaaa ccaacagatt aaaagccgct cgctattatg aatg 434
214 <210> SEQ ID NO: 6
215 <211> LENGTH: 110
216 <212> TYPE: PRT
217 <213> ORGANISM: Helicobacter pylori
219 <400> SEQUENCE: 6
220 Met Leu Lys Lys Ser Leu Leu Leu Leu Val Phe Leu Val Leu Gln Leu
221 1          5          10          15
222 Ser Gly Ala Glu Glu Asn Asn Gln Ala Pro Lys Asn Thr Pro Pro Glu
223          20          25          30
224 Leu Asn Pro Ala Asn Ala Lys Gly Ala Pro Asn Ser Asn Thr Gln Ile
225          35          40          45
226 Thr Pro Lys Asn Asp Asn Ser Asn Leu Leu Asp Lys Leu Gly Ser Pro
227          50          55          60
228 Glu Asn Ala Gln Thr Glu Leu Ser Ala Gly Ile Asp Leu Ala Lys Lys
229 65          70          75          80
230 Gly Asp Tyr Gln Gly Ala Phe Lys Leu Phe Ser Gln Ser Cys Asp Asn
231          85          90          95
232 Gly Asn Ala Ala Gly Cys Phe Ala Ser Gly Gly Asp Val Cys
233          100          105          110
236 <210> SEQ ID NO: 7
237 <211> LENGTH: 575
238 <212> TYPE: DNA
239 <213> ORGANISM: Helicobacter pylori
241 <220> FEATURE:
242 <221> NAME/KEY: CDS
243 <222> LOCATION: (73)...(522)
245 <400> SEQUENCE: 7
246 ccacaaaagc ttaatgaata ttaaattaaa aacatgttaa tctttagtta tttttaaaat 60
247 ttaggaaatc cc atg cat caa aac aat aaa act ttt tta ccc agc caa tcc 111

```

RAW SEQUENCE LISTING

DATE: 02/19/2002

PATENT APPLICATION: US/09/881,752A

TIME: 14:40:02

Input Set : A:\06132.041002.SEQLIST.TXT

Output Set: N:\CRF3\02192002\I881752A.raw

```

248           Met His Gln Asn Asn Lys Thr Phe Leu Pro Ser Gln Ser
249           1           5           10
251 gct cac ctc tct aaa atc att ctt ttt tta aac acc ggc ttt tta gcc 159
252 Ala His Leu Ser Lys Ile Ile Leu Phe Leu Asn Thr Gly Phe Leu Ala
253      15           20           25
255 tat ctg tta agc gct tgt ggg gcg aat gtg cct ata gaa gaa gtg ttg 207
256 Tyr Leu Leu Ser Ala Cys Gly Ala Asn Val Pro Ile Glu Glu Val Leu
257 30           35           40           45
259 gtt aaa gat cct aaa gag acc aaa gcc caa gaa gtc gcc aga gaa gaa 255
260 Val Lys Asp Pro Lys Glu Thr Lys Ala Gln Glu Val Ala Arg Glu Glu
261           50           55           60
263 aag gct atc cag caa gaa aac gcc act att gat gcg cgc acc acg cct 303
264 Lys Ala Ile Gln Gln Glu Asn Ala Thr Ile Asp Ala Arg Thr Thr Pro
265           65           70           75
267 tta atc aat cgt ttc act aat tat agc gct tat ggc tct tta aac ggc 351
268 Leu Ile Asn Arg Phe Thr Asn Tyr Ser Ala Tyr Gly Ser Leu Asn Gly
269      80           85           90
271 ttt tac aat tca gtg gat aat ctc aat tcg ccc atg caa aac ggg atg 399
272 Phe Tyr Asn Ser Val Asp Asn Leu Asn Ser Pro Met Gln Asn Gly Met
273 95           100           105
275 tat gga ggc tat tac atg cct tat tat tac atg ccc tat ggt ttc atg 447
276 Tyr Gly Gly Tyr Tyr Met Pro Tyr Tyr Tyr Met Pro Tyr Gly Phe Met
277 110           115           120           125
279 cct tat ggg tca ggt ctt atg cct tat ggg cct tat ggg tat gga gcg 495
280 Pro Tyr Gly Ser Gly Leu Met Pro Tyr Gly Pro Tyr Gly Tyr Gly Ala
281           130           135           140
283 cct gga tac ttc cct tac gct ttt tat tgattgagtg gcttttagaaa 542
284 Pro Gly Tyr Phe Pro Tyr Ala Phe Tyr
285           145           150
287 gcgtggtggt gttggtgttt ttactcaaac acg 575
289 <210> SEQ ID NO: 8
290 <211> LENGTH: 150
291 <212> TYPE: PRT
292 <213> ORGANISM: Helicobacter pylori
294 <400> SEQUENCE: 8
295 Met His Gln Asn Asn Lys Thr Phe Leu Pro Ser Gln Ser Ala His Leu
296 1           5           10           15
297 Ser Lys Ile Ile Leu Phe Leu Asn Thr Gly Phe Leu Ala Tyr Leu Leu
298      20           25           30
299 Ser Ala Cys Gly Ala Asn Val Pro Ile Glu Glu Val Leu Val Lys Asp
300 35           40           45
301 Pro Lys Glu Thr Lys Ala Gln Glu Val Ala Arg Glu Glu Lys Ala Ile
302 50           55           60
303 Gln Gln Glu Asn Ala Thr Ile Asp Ala Arg Thr Thr Pro Leu Ile Asn
304 65           70           75           80
305 Arg Phe Thr Asn Tyr Ser Ala Tyr Gly Ser Leu Asn Gly Phe Tyr Asn
306      85           90           95
307 Ser Val Asp Asn Leu Asn Ser Pro Met Gln Asn Gly Met Tyr Gly Gly
308 100           105           110

```

Use of n and/or Xaa has been detected in the Sequence Listing.
 Review the Sequence Listing to insure a corresponding
 explanation is presented in the <220> to <223> fields of
 each sequence using n or Xaa.

VERIFICATION SUMMARY

DATE: 02/19/2002

PATENT APPLICATION: US/09/881,752A

TIME: 14:40:03

Input Set : A:\06132.041002.SEQLIST.TXT

Output Set: N:\CRF3\02192002\I881752A.raw

L:1176 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19
L:1177 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19
L:1201 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:2050 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:29
L:2050 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29
L:2080 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:5159 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:77
L:5160 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:77
L:5208 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:78
L:5905 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:85
L:5906 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:85
L:5963 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:86
L:8082 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:125
L:8082 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:125
L:8106 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:126
L:8149 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:127
L:8149 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:127
L:8275 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:128
L:10939 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:153
L:14436 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:203
L:14541 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:204
L:17575 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:255
L:17576 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:255
L:17648 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:256
L:18984 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:275
L:19882 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:289
L:20255 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:299
L:20255 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:299
L:20327 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:300
L:20670 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:307
L:20670 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:307
L:21124 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:308
L:23178 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:337